1

PLANT GENE

TECHNICAL FIELD

The present invention relates to methods and materials for

5   controlling flowering time based on manipulation of
appropriate genes.

PRIOR ART

10   The timing of flowering in plants is of commercial interest
for a number of reasons. Particularly in those plants in
which the leaves or tubers are a commercial product, it is
desirable to avoid "bolting" (initiation of flowers and stem
elongation) at too early a stage. Conversely it may be

15   desirable to accelerate flowering under certain circumstances
e.g. to vary flower production across the seasons.
Horticultural plants whose flowering may be controlled
include lettuce, endive and vegetable brassicas including
cabbage, broccoli and cauliflower, and carnations and

20   geraniums.

It is known that flowering in plants may be dependent on a
number of external factors, which include photoperiodic
control and exposure to low temperatures. Certain plant

25   hormones and metabolites are known to be able to affect
flowering time. However the underlying mechanisms and
factors controlling the system are not fully understood.

Several studies of flowering have been carried out in

30   *Arabidopsis*. It has been shown that once the transition to
flowering happens in *Arabidopsis*, reversion to the vegetative
state rarely occurs. The cells of the apical meristem are
thus-considered stably determined for reproductive
development. Mutagenesis experiments have shown that the

35   transition and commitment to flowering in *Arabidopsis* is
controlled by many loci (Koornneef et al., 1998).

One locus–*FRIGIDA* (*FRI*), is responsible for the major
variation in flowering time in natural ecotypes of
Arabidopsis (Napp-Zinn, 1987; Lee et al., 1993; Burn et al.,
1993; Clarke and Dean, 1994; Sanda and Amasino, 1996). Late
5   flowering is conferred by dominant alleles at the *FRI* locus
in all crosses examined so far. The late flowering is very
extreme with plants producing around 80 rosette leaves before
starting to flower. This should be compared with 7 rosette
leaves in the early flowering ecotypes and about 25 in the
10 late flowering mutants at other loci. The extreme phenotype
of the late flowering ecotypes can be reversed by a
vernalization treatment where the imbibed seeds or plants are
incubated at low temperature (below 10°C) before growth at
normal temperature (20°C) (Napp-Zinn, 1965). Although the *FRI*
15 locus was not identified among the well characterized late
flowering mutants in the Landsberg *erecta* (L*er*) ecotype,
recent work has shown that this is due to recessive L*er*
alleles at a locus on chromosome 5 termed *FLC*. Dominant
alleles at *FLC* are required for the late flowering phenotype
20 conferred by *FRI* to be manifested (Lee et al., 1994;
Koornneef et al., 1994).

It is likely that *FRI* and *FLC* also play a major role in
determining whether Brassica plants require vernalization to
25 flower. Markers linked to the *FRI* and *FLC* have been shown to
co-segregate with the two QTL's conferring vernalization
requirement in Brassica species
(Osborn et al., 1997).

30 DISCLOSURE OF THE INVENTION

The present inventors have isolated a gene from the FRI locus
with a demonstrated effect in altering flowering times. The
gene shows only very little similarity with known sequences,
35 but appears to encode an alpha helical protein with no
membrane spanning domains. In various aspects of the
invention the gene, or its variants, may be employed in

manipulating the flowering time trait in a variety of plants.

According to a first aspect of the present invention there is
provided an isolated nucleic acid molecule from the FRI locus
5   of a plant encoding a polypeptide which is capable of
altering the flowering time of a plant into which the nucleic
acid is introduced.

The alteration in flowering time (which may produce a
10   correspondingly altered vegetative phase) may be assessed by
comparison with a plant in which the nucleic acid is not
present. It may be preferable to use a sample of plants in
each case. Flowering time may be measured directly, or
inferred from other factors e.g. significant change in leaf
15   numbers at flowering (see Koornneef et al, 1991).

Apart from a quantitative change in flowering characteristics
(with associated alteration in leaf numbers etc.) it is
preferred that characteristics unrelated to flowering in the
20   plant are substantially unchanged by the polypeptide, which
is to say that the polypeptide acts specifically on the
flowering characteristic.

Preferably the polypeptide is capable of delaying flowering
25   time (=extending a vegetative phase).

Nucleic acid according to the present invention may include
cDNA, RNA, genomic DNA and modified nucleic acids or nucleic
acid analogs (e.g. peptide nucleic acid). Where a DNA
30   sequence is specified, e.g. with reference to a figure,
unless context requires otherwise the RNA equivalent, with U
substituted for T where it occurs, is encompassed.
Likewise, as DNA is generally found in double-stranded form,
the complement of those sequences below is also included.
35   Nucleic acid molecules according to the present invention may
be provided isolated and/or purified from their natural
environment, in substantially pure or homogeneous form, or

free or substantially free of other nucleic acids of the
species of origin. Where used herein, the term "isolated"
encompasses all of these possibilities. The nucleic acid
molecules may be wholly or partially synthetic. In

5 particular they may be recombinant in that nucleic acid
sequences which are not found together in nature (do not run
contiguously) have been ligated or otherwise combined
artificially. Alternatively they may have been synthesised
directly e.g. using an automated synthesiser.

10

Most preferably the nucleic acid is derived from the FRI
locus of *Arabidopsis*.

15 Thus in one embodiment of this aspect of the invention, there
is disclosed a nucleic acid encoding the polypeptide of Fig
6.

A genomic sequence corresponding to the *Arabidopsis* FRI locus
is shown in Fig 4. A putative cDNA sequence transcribed from

20 this genomic sequence is shown at Fig 5. Preferably the
nucleic acid of the invention comprises the coding sequence
within Fig 5 (bases 362-2188 inclusive). More preferably the
nucleic acid comprises the sequence of Fig 5 or Fig 4.

25 In a further aspect of the present invention there are
disclosed nucleic acids which are variants of the sequences
provided.

A variant nucleic acid molecule shares homology with, or is

30 identical to, all or part of the coding sequence discussed
above. Generally variants may encode, or be used to isolate
or amplify nucleic acids which encode, polypeptides which are
capable of altering the flowering characteristics of plants
as described above. Variants of the present invention can be

35 artificial nucleic acids, which can be prepared by the
skilled person in the light of the present disclosure.
Alternatively they may be novel, naturally occurring, nucleic

acids, isolatable using the sequences of the present
invention.

Sequence variants which occur naturally may include FRI

5   alleles (which will include polymorphisms or mutations at one
or more bases) or pseudoalleles (which may occur at closely
linked loci to the FRI gene). Also included within the scope
of the present invention are isogenes, or other homologous
genes belonging to the same family as the FRI gene. Although

10   these may occur at different genomic loci to the gene, they
are likely to share conserved regions with it.

Certain allelic and homologous variants are discussed in the
Examples below.

15   Artificial variants (derivatives) may be prepared by those
skilled in the art, for instance by site directed or random
mutagenesis, or by direct synthesis. Preferably the variant
nucleic acid is generated either directly or indirectly (e.g.

20   via one or more amplification or replication steps) from an
original nucleic acid having all or part of the sequence
shown in Fig 5.

Thus a variant may be a distinctive part or fragment (however

25   produced) corresponding to a portion of the sequence
provided. The fragments may encode particular functional
parts of the polypeptide. Alternatively, the fragments may
have utility in probing for, or amplifying, the sequence
provided or closely related ones. Suitable lengths of

30   fragment, and conditions, for such processes are discussed in
more detail below.

Also included are nucleic acids corresponding to those above,
but which have been extended at the 3' or 5' terminus.

35

The term 'variant' nucleic acid as used herein encompasses
all of these possibilities. When used in the context of

6

polypeptides or proteins it indicates the encoded expression
product of the variant nucleic acid.

Some of the aspects of the present invention relating to
variants will now be discussed in more detail.

Similarity or homology (the terms are used interchangeably)
or identity may be as defined and determined by the TBLASTN
program, of Altschul *et al.* (1990) *J. Mol. Biol.* 215: 403-10,
or BestFit, which is part of the Wisconsin Package, Version
8, September 1994, (Genetics Computer Group, 575 Science
Drive, Madison, Wisconsin, USA, Wisconsin 53711). Preferably
sequence comparisons are made using FASTA and FASTP (see
Pearson & Lipman, 1988. Methods in Enzymology 183: 63-98).
Parameters are preferably set, using the default matrix, as
follows:

Gapopen (penalty for the first residue in a gap): -12 for
proteins / -16 for DNA
Gapext (penalty for additional residues in a gap): -2 for
proteins / -4 for DNA
KTUP word length: 2 for proteins / 6 for DNA.

Homology may be at the nucleotide sequence and/or encoded
amino acid sequence level. Preferably, the nucleic acid
and/or amino acid sequence shares at least about 60%, or 70%,
or 80% homology, most preferably at least about 90%, 95%,
96%, 97%, 98% or 99% homology or identity.

Homology may be over the full-length of the relevant sequence
shown herein, or may be over a part of it, preferably over a
contiguous sequence of about or greater than about 20, 25,
30, 33, 40, 50, 67, 133, 167, 200, 233, 267, 300, 333, 400 or
more amino acids or codons, compared with Fig 6 or 5
respectively.

Thus a variant polypeptide in accordance with the present

invention may include within the sequence shown in Fig 6, a
single amino acid or 2, 3, 4, 5, 6, 7, 8, or 9 changes, about
10, 15, 20, 30, 40 or 50 changes, or greater than about 50,
60, 70, 80 or 90 changes. In addition to one or more changes
5   within the amino acid sequence shown, a variant polypeptide
may include additional amino acids at the C-terminus and/or
N-terminus. Naturally, changes to the nucleic acid which make
no difference to the encoded polypeptide (i.e.
'degeneratively equivalent') are included.

10   The activity of functional variant polypeptides may be
assessed by transformation into a host capable of expressing
the nucleic acid of the invention. Methodology for such
transformation is described in more detail below.

15   In a further aspect of the invention there is disclosed a
method of producing a derivative nucleic acid comprising the
step of modifying any of the sequences disclosed above,
particularly the coding sequence of Fig 5.

20   Changes may be desirable for a number of reasons. For
instance they may introduce or remove restriction
endonuclease sites or alter codon usage.

25   Alternatively changes to a sequence may produce a derivative
by way of one or more of addition, insertion, deletion or
substitution of one or more nucleotides in the nucleic acid,
leading to the addition, insertion, deletion or substitution
of one or more amino acids in the encoded polypeptide.

30   Such changes may modify sites which are required for post
translation modification such as cleavage sites in the
encoded polypeptide; motifs in the encoded polypeptide for
glycosylation, lipoylation etc. Leader or other targeting
35   sequences (e.g. membrane or golgi locating sequences) may be
added to the expressed protein to determine its location
following expression.

8

Other desirable mutation may be random or site directed
mutagenesis in order to alter the activity (e.g. specificity)
or stability of the encoded polypeptide. Changes may be by
way of conservative variation, i.e. substitution of one
5   hydrophobic residue such as isoleucine, valine, leucine or
methionine for another, or the substitution of one polar
residue for another, such as arginine for lysine, glutamic
for aspartic acid, or glutamine for asparagine. As is well
known to those skilled in the art, altering the primary
10   structure of a polypeptide by a conservative substitution may
not significantly alter the activity of that peptide because
the side-chain of the amino acid which is inserted into the
sequence may be able to form similar bonds and contacts as
the side chain of the amino acid which has been substituted
15   out. This is so even when the substitution is in a region
which is critical in determining the peptides conformation.
Also included are variants having non-conservative
substitutions. As is well known to those skilled in the art,
substitutions to regions of a peptide which are not critical
20   in determining its conformation may not greatly affect its
activity because they do not greatly alter the peptide's
three dimensional structure. In regions which are critical
in determining the peptides conformation or activity such
changes may confer advantageous properties on the
25   polypeptide. Indeed, changes such as those described above
may confer slightly advantageous properties on the peptide
e.g. altered stability or specificity.

In a further aspect of the present invention there is
30   provided a method of identifying and/or cloning a nucleic
acid variant from a plant which method employs a FRI sequence
described above.

In one embodiment, nucleotide sequence information provided
35   herein may be used in a data-base (e.g. of ESTs, or STSs)
search to find homologous sequences, such as those which may
become available in due course, and expression products of

9

which can be tested for activity as described below.

In another embodiment the nucleotide sequence information
provided herein may be used to design probes and primers for
5   probing or amplification of FRI or variants thereof. An
oligonucleotide for use in probing or PCR may be about 30 or
fewer nucleotides in length (e.g. 18, 21 or 24). Generally
specific primers are upwards of 14 nucleotides in length.
For optimum specificity and cost effectiveness, primers of
10   16-24 nucleotides in length may be preferred. Those skilled
in the art are well versed in the design of primers for use
processes such as PCR. If required, probing can be done with
entire restriction fragments of the gene disclosed herein
which may be 100's or even 1000's of nucleotides in length.
15   Naturally sequences may be based on Fig 4 or Fig 5, or the
complement thereof. Small variations may be introduced into
the sequence to produce 'consensus' or 'degenerate' primers
if required.

20   Such probes and primers form one aspect of the present
invention.

Probing may employ the standard Southern blotting technique.
For instance DNA may be extracted from cells and digested
25   with different restriction enzymes. Restriction fragments
may then be separated by electrophoresis on an agarose gel,
before denaturation and transfer to a nitrocellulose filter.
Labelled probe may be hybridised to the DNA fragments on the
filter and binding determined. DNA for probing may be
30   prepared from RNA preparations from cells. Probing may
optionally be done by means of so-called 'nucleic acid chips'
(see Marshall & Hodgson (1998) Nature Biotechnology 16: 27-
31, for a review).

35   In one embodiment, a variant in accordance with the present
invention is obtainable by means of a method which includes:

(a) providing a preparation of nucleic acid, e.g. from plant cells. Test nucleic acid may be provided from a cell as genomic DNA, cDNA or RNA, or a mixture of any of these, preferably as a library in a suitable vector. If genomic DNA

5   is used the probe may be used to identify untranscribed regions of the gene (e.g. promoters etc.) as described hereinafter,

(b) providing a probe or primer as discussed above,

10

(c) contacting nucleic acid in said preparation with said nucleic acid molecule under conditions for hybridisation of said nucleic acid molecule to any said gene or homologue in said preparation, and,

15

(d) identifying said gene or homologue if present by its hybridisation with said nucleic acid molecule. Binding of a probe to target nucleic acid (e.g. DNA) may be measured using any of a variety of techniques at the disposal of those

20   skilled in the art. For instance, probes may be radioactively, fluorescently or enzymatically labelled. Other methods not employing labelling of probe include amplification using PCR (see below), RN'ase cleavage and allele specific oligonucleotide probing. The identification

25   of successful hybridisation is followed by isolation of the nucleic acid which has hybridised, which may involve one or more steps of PCR or amplification of a vector in a suitable host.

30   Preliminary experiments may be performed by hybridising under low stringency conditions. For probing, preferred conditions are those which are stringent enough for there to be a simple pattern with a small number of hybridisations identified as positive which can be investigated further.

35

For example, hybridizations may be performed, according to the method of Sambrook et al. (below) using a hybridization

solution comprising: 5X SSC (wherein 'SSC' = 0.15 M sodium
chloride; 0.15 M sodium citrate; pH 7), 5X Denhardt's
reagent, 0.5-1.0% SDS, 100 µg/ml denatured, fragmented salmon
sperm DNA, 0.05% sodium pyrophosphate and up to 50%

5   formamide. Hybridization is carried out at 37-42°C for at
least six hours. Following hybridization, filters are washed
as follows: (1) 5 minutes at room temperature in 2X SSC and
1% SDS; (2) 15 minutes at room temperature in 2X SSC and 0.1%
SDS; (3) 30 minutes - 1 hour at 37°C in 1X SSC and 1% SDS; (4)

10 2 hours at 42-65°C in 1X SSC and 1% SDS, changing the solution
every 30 minutes.

One common formula for calculating the stringency conditions
required to achieve hybridization between nucleic acid

15 molecules of a specified sequence homology is (Sambrook et
al., 1989): T_m = 81.5°C + 16.6Log [Na+] + 0.41 (% G+C) - 0.63
(% formamide) - 600/#bp in duplex

As an illustration of the above formula, using [Na+] =

20 [0.368] and 50-% formamide, with GC content of 42% and an
average probe size of 200 bases, the T_m is 57°C. The T_m of a
DNA duplex decreases by 1 - 1.5°C with every 1% decrease in
homology. Thus, targets with greater than about 75% sequence
identity would be observed using a hybridization temperature

25 of 42°C. Such a sequence would be considered substantially
homologous to the nucleic acid sequence of the present
invention.

It is well known in the art to increase stringency of

30 hybridisation gradually until only a few positive clones
remain. Other suitable conditions include, e.g. for
detection of sequences that are about 80-90% identical,
hybridization overnight at 42°C in 0.25M Na_2HPO_4, pH 7.2, 6.5%
SDS, 10% dextran sulfate and a final wash at 55°C in 0.1X

35 SSC, 0.1% SDS. For detection of sequences that are greater
than about 90% identical, suitable conditions include
hybridization overnight at 65°C in 0.25M Na_2HPO_4, pH 7.2, 6.5%

12

SDS, 10% dextran sulfate and a final wash at 60°C in 0.1X
SSC, 0.1% SDS.

An alternative, which may be particularly appropriate with
5   plant nucleic acid preparations, is a solution of 5x SSPE
(final 0.9 M NaCl, 0.05M sodium phosphate, 0.005M EDTA pH
7.7), 5X Denhardt's solution, 0.5% SDS, at 65°C overnight,
(for high stringency, highly similar sequences) or 50°C (for
low stringency, less similar sequences). Washes in 0.2x
10   SSC/0.1% SDS at 65°C for high stringency, alternatively at
50-60°C in 1x SSC/0.1%SDS for low stringency.

In a further embodiment, hybridisation of nucleic acid
molecule to a variant may be determined or identified
15   indirectly, e.g. using a nucleic acid amplification reaction,
particularly the polymerase chain reaction (PCR). PCR
requires the use of two primers to specifically amplify
target nucleic acid, so preferably two nucleic acid molecules
with sequences characteristic of FRI are employed. Using RACE
20   PCR, only one such primer may be needed (see "PCR protocols;
A Guide to Methods and Applications", Eds. Innis et al,
Academic Press, New York, (1990)).

Thus a method involving use of PCR in obtaining nucleic acid
25   according to the present invention may include:

(a) providing a preparation of plant nucleic acid, e.g. from
a plant cell,

30   (b) providing a pair of nucleic acid molecule primers useful
in (i.e. suitable for) PCR, at least one of said primers
being a primer according to the present invention as
discussed above,

35   (c) contacting nucleic acid in said preparation with said
primers under conditions for performance of PCR,

(d) performing PCR and determining the presence or absence of
an amplified PCR product. The presence of an amplified PCR
product may indicate identification of a variant.

5   In all cases above, if need be, clones or fragments
identified in the search can be extended. For instance if it
is suspected that they are incomplete, the original DNA
source (e.g. a clone library, mRNA preparation etc.) can be
revisited to isolate missing portions e.g. using sequences,
10   probes or primers based on that portion which has already
been obtained to identify other clones containing overlapping
sequence. As used hereinafter, unless the context demands
otherwise, the term "FRI" is intended to cover any of the
nucleic acids of the invention described above, including
15   functional variants.

The methods described above may also be used to determine the
presence of one of the nucleotide sequences of the present
invention within the genetic context of an individual plant,
20   optionally a transgenic plant such as may be produced as
described in more detail below. This may be useful in plant
breeding programmes e.g. to directly select plants containing
alleles which are responsible for desirable traits in that
plant species, either in parent plants or in progeny (e.g
25   hybrids, F1, F2 etc.). Thus use of the newly defined markers
disclosed in the Examples below, or markers which can be
designed by those skilled in the art on the basis of the
nucleotide sequence information disclosed herein, for
selection of a gene capable of modifying flowering time in a
30   plant, forms one part of the present invention.

The use of diagnostic tests for alleles allows the researcher
or plant breeder to establish, with full confidence and
independent from time consuming tests based on actual
35   observation of desired traits (in this case, flowering
properties), whether or not a desired allele is present in
the plant of interest (or a cell thereof), whether the plant

is a representative of a collection of other genetically
identical plants (e.g. an inbred variety or cultivar) or one
individual in a sample of related (e.g. breeders' selection)
or unrelated plants. In a breeding scheme based on selection

5   and selfing of desirable individuals, nucleic acid or
polypeptide diagnostics for the desirable allele or alleles
in high throughput, low cost assays as provided by this
invention, reliable selection for the flowering time
phenotype can be made at early generations and on more

10   material than would otherwise be possible. This gain in
reliability of selection plus the time saving by being able
to test material earlier and without costly phenotype
screening is of considerable value in plant breeding.

15   Nucleic acid-based determination of the presence or absence
of one or more desirable alleles may be combined with
determination of the genotype of the flanking linked genomic
DNA and other unlinked genomic DNA using markers such as
RFLPs, microsatellites or SSRs, AFLPs, RAPDs etc which are

20   either very closely linked to FRI locus, as shown in the
Examples hereinafter, or are adapted to identify individual
FRI alleles for direct allele selection. This enables the
researcher or plant breeder to select for not only the
presence of the desirable allele but also for individual

25   plant or families of plants which have the most desirable
combinations of linked and unlinked genetic background. Such
recombinations of desirable material may occur only rarely
within a given segregating breeding population or backcross
progeny. Direct assay of the locus as afforded by the

30   present invention allows the researcher to make a step-wise
approach to fixing (making homozygous) the desired
combination of flanking markers and alleles, by first
identifying individuals fixed for one flanking marker and
then identifying progeny fixed on the other side of the locus

35   all the time knowing with confidence that the desirable
allele is still present.

In a further aspect of the present invention, nucleic acid encoding FRI is in the form of a recombinant and preferably replicable vector.

5   "Vector" is defined to include, inter alia, any plasmid, cosmid, phage or *Agrobacterium* binary vector in double or single stranded linear or circular form which may or may not be self transmissible or mobilizable, and which can transform a prokaryotic or eukaryotic host either by integration into

10   the cellular genome or exist extrachromosomally (e.g. autonomous replicating plasmid with an origin of replication).

Generally speaking, those skilled in the art are well able to
15   construct vectors and design protocols for recombinant gene expression. Suitable vectors can be chosen or constructed, containing appropriate regulatory sequences, including promoter sequences, terminator fragments, polyadenylation sequences, enhancer sequences, marker genes and other
20   sequences as appropriate. For further details see, for example, *Molecular Cloning: a Laboratory Manual:* 2nd edition, Sambrook *et al,* 1989, Cold Spring Harbor Laboratory Press or *Current Protocols in Molecular Biology,* Second Edition, Ausubel et al. eds., John Wiley & Sons, 1992.

25
Specifically included are shuttle vectors by which is meant a DNA vehicle capable, naturally or by design, of replication in two different host organisms, which may be selected from actinomycetes and related species, bacteria and eucaryotic
30   (e.g. higher plant, mammalian, yeast or fungal cells).

A vector including nucleic acid according to the present invention need not include a promoter or other regulatory sequence, particularly if the vector is to be used to
35   introduce the nucleic acid into cells for recombination into the genome.

16

Preferably the nucleic acid in the vector is under the
control of, and operably linked to, an appropriate promoter
or other regulatory elements for transcription in a host cell
such as a microbial, e.g. bacterial, or plant cell. The
5   vector may be a bi-functional expression vector which
functions in multiple hosts. In the case of FRI genomic DNA,
this may contain its own promoter or other regulatory
elements and in the case of cDNA this may be under the
control of an appropriate promoter or other regulatory
10   elements for expression in the host cell.

By "promoter" is meant a sequence of nucleotides from which
transcription may be initiated of DNA operably linked
downstream (i.e. in the 3' direction on the sense strand of
15   double-stranded DNA).

"Operably linked" means joined as part of the same nucleic
acid molecule, suitably positioned and oriented for
transcription to be initiated from the promoter. DNA
20   operably linked to a promoter is "under transcriptional
initiation regulation" of the promoter.

In a preferred embodiment, the promoter is an inducible
promoter.
25

The term "inducible" as applied to a promoter is well
understood by those skilled in the art. In essence,
expression under the control of an inducible promoter is
"switched on" or increased in response to an applied
30   stimulus. The nature of the stimulus varies between
promoters. Some inducible promoters cause little or
undetectable levels of expression (or no expression) in the
absence of the appropriate stimulus. Other inducible
promoters cause detectable constitutive expression in the
35   absence of the stimulus. Whatever the level of expression is
in the absence of the stimulus, expression from any inducible
promoter is increased in the presence of the correct

17

stimulus.

Thus this aspect of the invention provides a gene construct,
preferably a replicable vector, comprising a promoter
5   (optionally inducible) operably linked to a nucleotide
sequence provided by the present invention, such as the FRI
gene or a variant thereof.

Particular of interest in the present context are nucleic
10   acid constructs which operate as plant vectors. Specific
procedures and vectors previously used with wide success upon
plants are described by Guerineau and Mullineaux (1993)
(Plant transformation and expression vectors. In: Plant
Molecular Biology Labfax (Croy RRD ed) Oxford, BIOS
15   Scientific Publishers, pp 121-148). Suitable vectors may
include plant viral-derived vectors (see e.g. EP-A-194809).

Suitable promoters which operate in plants include the
Cauliflower Mosaic Virus 35S (CaMV 35S). Other examples are
20   disclosed at pg 120 of Lindsey & Jones (1989) "Plant
Biotechnology in Agriculture" Pub. OU Press, Milton Keynes,
UK. The promoter may be selected to include one or more
sequence motifs or elements conferring developmental and/or
tissue-specific regulatory control of expression. Inducible
25   plant promoters include the ethanol induced promoter of
Caddick et al (1998) Nature Biotechnology 16: 177-180.

If desired, selectable genetic markers may be included in the
construct, such as those that confer selectable phenotypes
30   such as resistance to antibiotics or herbicides (e.g.
kanamycin, hygromycin, phosphinotricin, chlorsulfuron,
methotrexate, gentamycin, spectinomycin, imidazolinones and
glyphosate).

35   The present invention also provides methods comprising
introduction of such a construct into a plant cell or a
microbial cell and/or induction of expression of a construct

within a plant cell, by application of a suitable stimulus
e.g. an effective exogenous inducer.

The invention further encompasses a host cell transformed
5   with nucleic acid or a vector according to the present
invention (e.g comprising the FRI sequence) especially a
plant or a microbial cell. In the transgenic plant cell
(i.e. transgenic for the nucleic acid in question) the
transgene may be on an extra-genomic vector or incorporated,
10   preferably stably, into the genome. There may be more than
one heterologous nucleotide sequence per haploid genome.

The term "heterologous" is used broadly in this aspect to
indicate that the gene/sequence of nucleotides in question
15   (e.g. encoding FRI) have been introduced into said cells of
the plant or an ancestor thereof, using genetic engineering,
i.e. by human intervention. A heterologous gene may replace
an endogenous equivalent gene, i.e. one which normally
performs the same or a similar function, or the inserted
20   sequence may be additional to the endogenous gene or other
sequence. Nucleic acid heterologous to a plant cell may be
non-naturally occurring in cells of that type, variety or
species. Thus the heterologous nucleic acid may comprise a
coding sequence of or derived from a particular type of plant
25   cell or species or variety of plant, placed within the
context of a plant cell of a different type or species or
variety of plant. A further possibility is for a nucleic
acid sequence to be placed within a cell in which it or a
homologue is found naturally, but wherein the nucleic acid
30   sequence is linked and/or adjacent to nucleic acid which does
not occur naturally within the cell, or cells of that type or
species or variety of plant, such as operably linked to one
or more regulatory sequences, such as a promoter sequence,
for control of expression.

35

Thus a further aspect of the present invention provides a
method of transforming a plant cell involving introduction of

19

a construct as described above into a plant cell and causing
or allowing recombination between the vector and the plant
cell genome to introduce a nucleic acid according to the
present invention into the genome.

5

The host cell (e.g. plant cell) is preferably transformed by
the construct, which is to say that the construct becomes
established within the cell, altering one or more of the
cell's characteristics.

10

Nucleic acid can be introduced into plant cells using any
suitable technology, such as a disarmed Ti-plasmid vector
carried by Agrobacterium exploiting its natural gene transfer
ability (EP-A-270355, EP-A-0116718, NAR 12(22) 8711 - 87215

15   1984), particle or microprojectile bombardment (US 5100792,
EP-A-444882, EP-A-434616) microinjection (WO 92/09696, WO
94/00583, EP 331083, EP 175966, Green et al. (1987) Plant
Tissue and Cell Culture, Academic Press), electroporation (EP
290395, WO 8706614 Gelvin Debeyser) other forms of direct DNA

20   uptake (DE 4005152, WO 9012096, US 4684611), liposome
mediated DNA uptake (e.g. Freeman et al. Plant Cell Physiol.
29: 1353 (1984)), or the vortexing method (e.g. Kindle, PNAS
U.S.A. 87: 1228 (1990d) Physical methods for the
transformation of plant cells are reviewed in Oard, 1991,

25   Biotech. Adv.   9: 1-11.

Agrobacterium transformation is widely used by those skilled
in the art to transform dicotyledonous species.

30   Recently, there has also been substantial progress towards
the routine production of stable, fertile transgenic plants
in almost all economically relevant monocot plants (see e.g.
Hiei et al. (1994) The Plant Journal 6, 271-282)).
Microprojectile bombardment, electroporation and direct DNA

35   uptake are preferred where Agrobacterium alone is inefficient
or ineffective. Alternatively, a combination of different
techniques may be employed to enhance the efficiency of the

transformation process, eg bombardment with Agrobacterium
coated microparticles (EP-A-486234) or microprojectile
bombardment to induce wounding followed by co-cultivation
with Agrobacterium (EP-A-486233).

The particular choice of a transformation technology will be
determined by its efficiency to transform certain plant
species as well as the experience and preference of the
person practising the invention with a particular methodology
of choice. It will be apparent to the skilled person that
the particular choice of a transformation system to introduce
nucleic acid into plant cells is not essential to or a
limitation of the invention.

Generally speaking, following transformation, a plant may be
regenerated, e.g. from single cells, callus tissue or leaf
discs, as is standard in the art. Almost any plant can be
entirely regenerated from cells, tissues and organs of the
plant. Available techniques are reviewed in Vasil et al.,
*Cell Culture and Somatic Cell Genetics of Plants, Vol I, II
and III, Laboratory Procedures and Their Applications*,
Academic Press, 1984, and Weissbach and Weissbach, *Methods
for Plant Molecular Biology*, Academic Press, 1989.

The generation of fertile transgenic plants has been achieved
in the cereals rice, maize, wheat, oat, and barley (reviewed
in Shimamoto, K. (1994) *Current Opinion in Biotechnology* 5,
158-162.; Vasil, *et al.* (1992) *Bio/Technology* 10, 667-674;
Vain et al., 1995, *Biotechnology Advances* 13 (4): 653-671;
Vasil, 1996, *Nature Biotechnology* 14 page 702).

Plants which include a plant cell according to the invention
are also provided.

In addition to the regenerated plant, the present invention
embraces all of the following: a clone of such a plant, seed,
selfed or hybrid progeny and descendants (e.g. F1 and F2

descendants).

A plant according to the present invention may be one which
does not breed true in one or more properties. Plant
5   varieties may be excluded, particularly registrable plant
varieties according to Plant Breeders' Rights. It is noted
that a plant need not be considered a "plant variety" simply
because it contains stably within its genome a transgene,
introduced into a cell of the plant or an ancestor thereof.

10

The invention also provides a plant propagule from such
plants, that is any part which may be used in reproduction or
propagation, sexual or asexual, including cuttings, seed and
so on. It also provides any part of these plants e.g. edible
15   leaves which includes the plant cells or heterologous DNA
described above.

Preferred plants which may be transformed with FRI (thereby
delaying flowering time) include sugar beet, Brassicas
20   (especially cauliflower, broccoli, cabbage, spinach, curly
kale, *B. Napus*) and potato, lettuce and culinary herbs.

It may be preferable to use FRI in conjunction with other
genes or mutations which affect the flowering time of plants.
25   Examples include use of the FLC gene, which may enhance the
effects the FRI gene.

In particular it may be preferred to employ FRI in plants
which exhibit a reduced vernalization response. One means by
30   which this may be achieved is by means of the VRN2 gene.
VRN2 is believed to mediate the vernalisation response in
plants. Down-regulation of VRN2 expression, for instance by
means of an antisense VRN2 cDNA, may therefore be used to
inhibit the effect of vernalisation in promoting flowering.
35   VRN2 cDNA sequences from *Arabidopsis thaliana* Landsberg
erecta and Columbia are shown after the References. Anti-
sense methodology is discussed below.

22

The invention further provides a method of influencing or affecting a flowering time in a plant, the method including the step of causing or allowing expression of a heterologous FRI nucleic acid sequence as discussed above within the cells of the plant.

The invention further provides a method comprising the step of causing or allowing expression of a nucleic acid encoding FRI or a variant thereof, within cells of a plant (thereby producing the encoded polypeptide) such as to alter the flowering time.

The step may be preceded by the earlier step of introduction of the nucleic acid into a cell of the plant or an ancestor thereof.

The foregoing discussion has been generally concerned with uses of the nucleic acids of the present invention for production of functional polypeptides, thereby increasing the flowering time of the plant.

However the information disclosed herein may also be used to reduce the activity of FRI in cells in which it is desired to do so, thereby having the opposite effect.

Accelerating flowering time may be useful in certain species e.g. soft fruits such as strawberries, raspberries, or crops such as maize.

Down-regulation of FRI expression may, for instance, be achieved using anti-sense technology.

In using anti-sense genes or partial gene sequences to down-regulate gene expression, a nucleotide sequence is placed under the control of a promoter in a "reverse orientation" such that transcription yields RNA which is complementary to normal mRNA transcribed from the "sense"

strand of the target gene. See, for example, Rothstein et
al, 1987; Smith *et al*, (1988) *Nature* 334, 724-726; Zhang *et
al*, (1992) *The Plant Cell* 4, 1575-1588, English *et al*., (1996)
The Plant Cell 8, 179-188. Antisense technology is also
reviewed in Bourque, (1995), *Plant Science* 105, 125-149, and
Flavell, (1994) *PNAS USA* 91, 3490-3496.

Thus a nucleotide sequence which is complementary to any of
those coding sequences disclosed above forms one part of the
present invention.

An alternative to anti-sense is to use a copy of all or part
of the target gene inserted in sense, that is the same,
orientation as the target gene, to achieve reduction in
expression of the target gene by co-suppression. See, for
example, van der Krol *et al*., (1990) *The Plant Cell* 2, 291-
299; Napoli *et al*., (1990) *The Plant Cell* 2, 279-289; Zhang
et al., (1992) *The Plant Cell* 4, 1575-1588, and US-A-
5,231,020. Further refinements of the gene silencing or co-
suppression technology may be found in WO95/34668
(Biosource); Angell & Baulcombe (1997) The EMBO Journal
16,12:3675-3684; and Voinnet & Baulcombe (1997) Nature 389:
pg 553.

Further options for down regulation of gene expression
include the use of ribozymes, e.g. hammerhead ribozymes,
which can catalyse the site-specific cleavage of RNA, such as
mRNA (see e.g. Jaeger (1997) "The new world of ribozymes"
Curr Opin Struct Biol 7:324-335, or Gibson & Shillitoe
(1997)"Ribozymes: their functions and strategies form their
use" Mol Biotechnol 7: 242-251.)

The complete sequence corresponding to the coding sequence
(in reverse orientation for anti-sense) need not be used.
For example fragments of sufficient length may be used. It
is a routine matter for the person skilled in the art to
screen fragments of various sizes and from various parts of

24

the coding sequence to optimise the level of anti-sense
inhibition. It may be advantageous to include the initiating
methionine ATG codon, and perhaps one or more nucleotides
upstream of the initiating codon. A further possibility is
5   to target a conserved sequence of a gene, e.g. a sequence
that is characteristic of one or more genes, such as a
regulatory sequence.

The sequence employed may be about 500 nucleotides or less,
10 possibly about 400 nucleotides, about 300 nucleotides, about
200 nucleotides, or about 100 nucleotides. It may be
possible to use oligonucleotides of much shorter lengths, 14-
23 nucleotides, although longer fragments, and generally even
longer than about 500 nucleotides are preferable where
15 possible, such as longer than about 600 nucleotides, than
about 700 nucleotides, than about 800 nucleotides, than about
1000 nucleotides or more.

It may be preferable that there is complete sequence identity
20 in the sequence used for down-regulation of expression of a
target sequence, and the target sequence, although total
complementarity or similarity of sequence is not essential.
One or more nucleotides may differ in the sequence used from
the target gene. Thus, a sequence employed in a down-
25 regulation of gene expression in accordance with the present
invention may be a wild-type sequence (e.g. gene) selected
from those available, or a variant of such a sequence.

The sequence need not include an open reading frame or
30 specify an RNA that would be translatable. It may be
preferred for there to be sufficient homology for the
respective anti-sense and sense RNA molecules to hybridise.
There may be down regulation of gene expression even where
there is about 5%, 10%, 15% or 20% or more mismatch between
35 the sequence used and the target gene. Effectively, the
homology should be sufficient for the down-regulation of gene
expression to take place.

Thus the present invention further provides the use of the
nucleotide sequence of Fig 5, or its complement, or a variant
of either for down-regulation of gene expression,
particularly down-regulation of expression of the FRI gene or

5   a homologue thereof, preferably in order to influence
(accelerate) the flowering time of a plant.

Anti-sense or sense regulation may itself be regulated by
employing an inducible promoter in an appropriate construct.

10

Although the above description has generally been concerned
with the transcribed parts of the FRI gene (in sense or
antisense)and variants and products thereof, also embraced
within the present invention are untranscribed parts of the

15   gene. Thus a further aspect of the invention is a nucleic
acid molecule encoding the promoter or other transcriptional
control sequence of the FRI gene.

The promoter region or other control sequences may be readily

20   identified on the basis of the genomic sequence shown in Fig
4 using a probe or primer as described above in relation to
the isolation of variants. Generally they will be found 5'
to the open reading frame of the gene and are obtainable by
probing a genomic DNA library with a nucleic acid of the

25   invention, selecting a clone which hybridizes under
conditions of medium to high stringency, and sequencing the
clone 5' to the open reading frame of the gene. Where only a
small amount of sequence is present in the 5' region, this
sequence may be used to reprobe the library to genome walk

30   further upstream. Analysis of the upstream region will
reveal control regions for gene expression including control
regions common to many genes (i.e TATA and CAAT boxes) and
other control regions, usually located from 1 to 10,000, such
as 1 to 1000 or 50 to 500 nucleotides upstream of the start

35   of transcription. Sequences identified as described above
can be assessed for promoter activity.

26

"Promoter activity" is used to refer to ability to initiate
transcription. The level of promoter activity is
quantifiable for instance by assessment of the amount of mRNA
produced by transcription from the promoter or by assessment
5   of the amount of protein product produced by translation of
mRNA produced by transcription from the promoter. The amount
of a specific mRNA present in an expression system may be
determined for example using specific oligonucleotides which
are able to hybridise with the mRNA and which are labelled or
10   may be used in a specific amplification reaction such as the
polymerase chain reaction.

Use of a reporter gene facilitates determination of promoter
activity by reference to protein production. The reporter
15   gene preferably encodes an enzyme which catalyses a reaction
which produces a detectable signal, preferably a visually
detectable signal, such as a coloured product. Many examples
are known, including β-galactosidase and luciferase.

20   Those skilled in the art are well aware of a multitude of
possible reporter genes and assay techniques which may be
used to determine promoter activity. Any suitable
reporter/assay may be used and it should be appreciated that
no particular choice is essential to or a limitation of the
25   present invention.

Thus in one aspect of the invention there is provided a
nucleic acid construct, preferably an expression vector,
including the FRI promoter region or fragment operably linked
30   to a heterologous gene, e.g. a coding sequence, which is
preferably not the coding sequence with which the promoter is
operably linked in nature.

The present invention also encompasses the expression product
35   of any of the coding (sense) nucleic acid sequences disclosed
above, and methods of making the expression product by
expression from encoding nucleic acid therefore under

suitable conditions, which may be in suitable host cells.

Following expression, the product may be isolated from the
expression system (e.g. microbial) and may be used as
5   desired.

The present invention particularly provides for the
production and use of fragments of the full-length
polypeptides disclosed herein, especially active portions
10   thereof. An "active portion" of a polypeptide means a
peptide which is less than said full length polypeptide, but
which retains an essential biological activity. In
particular, the active portion retains the ability to alter
flowering time in a plant such as *Arabidopsis thaliana*.
15

A "fragment" of a polypeptide means a stretch of amino acid
residues of at least about five to seven contiguous amino
acids, often at least about seven to nine contiguous amino
acids, typically at least about nine to 13 contiguous amino
20   acids and, most preferably, at least about 20 to 30 or more
contiguous amino acids. Fragments of the polypeptides may
include one or more epitopes useful for raising antibodies to
a portion of any of the amino acid sequences disclosed
herein. Preferred epitopes are those to which antibodies are
25   able to bind specifically, which may be taken to be binding a
polypeptide or fragment thereof of the invention with an
affinity which is at least about 1000x that of other
polypeptides.

30   For instance, purified FRI protein, or a variant thereof,
e.g. produced recombinantly by expression from encoding
nucleic acid therefor, may be used to raise antibodies
employing techniques which are standard in the art.
Antibodies and polypeptides comprising antigen-binding
35   fragments of antibodies may be used in identifying homologues
from other species as discussed further below.

Methods of producing antibodies include immunising a mammal
(e.g. human, mouse, rat, rabbit, horse, goat, sheep or
monkey) with the protein or a fragment thereof. Antibodies
may be obtained from immunised animals using any of a variety
5   of techniques known in the art, and might be screened,
preferably using binding of antibody to antigen of interest.

For instance, Western blotting techniques or
immunoprecipitation may be used (Armitage et al, 1992, Nature
10   357: 80-82).

Antibodies may be polyclonal or monoclonal.

Antibodies may be modified in a number of ways. Indeed the
15   term "antibody" should be construed as covering any specific
binding substance having a binding domain with the required
specificity. Thus, this term covers antibody fragments,
derivatives, functional equivalents and homologues of
antibodies, including any polypeptide comprising an
20   immunoglobulin binding domain, whether natural or synthetic.
Chimaeric molecules comprising an immunoglobulin binding
domain, or equivalent, fused to another polypeptide are
therefore included. Cloning and expression of Chimaeric
antibodies are described in EP-A-0120694 and EP-A-0125023.
25   It has been shown that fragments of a whole antibody can
perform the function of binding antigens. Examples of
binding fragments are well known to those skilled in the art.

As an alternative or supplement to immunising a mammal,
30   antibodies with appropriate binding specificity may be
obtained from a recombinantly produced library of expressed
immunoglobulin variable domains, e.g. using lambda
bacteriophage or filamentous bacteriophage which display
functional immunoglobulin binding domains on their surfaces;
35   for instance see WO92/01047.

Antibodies raised to a polypeptide or peptide can be used in

the identification and/or isolation of variant polypeptides,
and then their encoding genes. Thus, the present invention
provides a method of identifying or isolating a FRI
polypeptide or variant thereof (as discussed above),
comprising screening candidate polypeptides with a
polypeptide comprising the antigen-binding domain of an
antibody (for example whole antibody or a fragment thereof)
which is able to bind said FRI polypeptide or variant
thereof, or preferably has binding specificity for such a
polypeptide.

Specific binding members such as antibodies and polypeptides
comprising antigen binding domains of antibodies that bind
and are preferably specific for a FRI polypeptide or mutant
or derivative thereof represent further aspects of the
present invention, as do their use and methods which employ
them.

The invention will now be further described with reference to
the following non-limiting Figures and Examples. Other
embodiments of the invention will occur to those skilled in
the art in the light of these.

FIGURES

Fig 1 shows a physical map of the genomic region covering the
FRI from the Arabidopsis Columbia ecotype based on a YAC/BAC
contig.

Fig 2 shows a cosmid contig covering FRI, with the
complementing cosmid (84M13) shown shaded.

Fig 3 shows various subclones of 84M13 which were used to
narrow down the *FRI* region within the cosmid.

Fig 4 shows the (late flowering) H51 genomic sequence of the
first 17 kb of 84M13 encompassing the FRI gene.

Fig 5 shows the likely cDNA sequence of the H51 FRI gene.

Fig 6 shows the FRI amino acid sequence predicted from the cDNA sequence.

5

Fig 7 shows Table 3.

EXAMPLES

10 Example 1- Mapping and isolation of the FRIGIDA gene

Genetic fine mapping of the FRI locus.

Co-segregation analysis of *FRI* with markers mi51, mi122,
15   mi204 and g3843 had mapped the *FRI* locus near the top of
chromosome 4 (Clarke and Dean, 1994). These and additional
markers, derived from the physical map of chromosome 4
(Schmidt et al., 1995) were then used to fine-map the locus.
590 F2 plants from a cross between the late flowering parent
20   H51 (Napp-Zinn, 1957; Napp-Zinn, 1962) and the early
flowering parent Li5 were scored for their flowering time.
344 of the plants flowered at the same time as H51 (late) and
146 plants flowered at the same time as Li-5 (early). This
3:1 ratio of late to early flowering time confirmed that *FRI*
25   segregates as a single locus with late flowering being the
dominant phenotype. The early flowering F2 plants were
allowed to self. The resulting F3 plants were scored for
their flowering time to confirm that they were homozygous for
the Li-5 allele at *FRI*. The segregation of the four markers
30   described above, closely linked to *FRI*, was then scored in
the 146 F2 early flowering plants. DNA from the F3 families
was digested with the restriction endonucleases that gave
polymorphisms between H51 and Li-5 and Southern blots were
prepared. These were then probed with the genetic markers.
35   Since all the F3 plants had the Li-5 allele at the *FRI* locus
and the DNA markers had been found to be either closely or
tightly linked to *FRI* it was anticipated that most of the F3

plants would show the Li-5 hybridisation patterns on the
resulting autoradiographs. However, in a few of the plants
there may have been a recombination event very close to *FRI* .
These plants would then show a heterozygous pattern with one
5   or more of the DNA markers. Two recombinant plants were
found in the population of 146 early flowering plants. Plant
number 156 (p156) showed a heterozygous pattern with marker
mi51, but a Li5 pattern with marker g8802, mi204, mi122 and
g3843. Plant number 5 (p5) showed a heterozygous pattern with
10 marker g3843 but a Li5 pattern with marker mi122, mi204 and
mi51. These data placed *FRI* between mi51 and g3843 on
chromosome 4, a genomic region of 350kb.

A further 180 early flowering F2 plants from an H51 X
Li-5 cross were screened with both mi51 and g3843 (both
15 markers showed a polymorphism between H51 and Li-5 with the
restriction endonuclease *Dra* I). In addition, 78 plants were
scored with mi51 only and 29 plants were scored with g3843
only. A total of five recombinant plants resulted from this
screen; three plants showed a heterozygous pattern with mi51
20 and two plants showed a heterozygous pattern with g3843.
These recombinant plants were checked to see if they gave a
heterozygous pattern with other markers in the region and
between which interval the recombination event had occurred.
One that was heterozygous for g3843 was also heterozygous for
25 CC27P11 so defining the location of FRI to an interval of
320kb.

Physical mapping of FRI.

30 The 320 kb region on the top of chromosome 4 between RFLP
markers mi51 and CC27P11 was covered by a physical map
comprising of yeast artificial chromosome (YAC) clones
(Clarke and Dean, 1994). Bacterial artificial chromosome
clones (BACs) were integrated into this map through the
35 efforts of scientists at Cold Spring Harbor Laboratory USA
(http://www.cshl.org/protarab/chrom4map.htm) and Washington
University USA

32

(http://genome.wustl.edu/gsc/arab/arabidopsis.html). The
physical map of the genomic region covering *FRI* from the
Arabidopsis Columbia ecotype is shown in Fig.1

5   *Complementation and further mapping.*

The genetics of the flowering time suggested that late
flowering ecotypes contained functional *FRI* alleles whereas
early flowering ecotypes contained only partially functional

10   or non-functional alleles. The libraries available for
complementation experiments had all been generated from early
flowering ecotypes (Landsberg *erecta*, Columbia and WS). In
order to do complementation experiments for *FRI*, a cosmid
library had to be made from an ecotype carrying a dominant

15   *FRI* allele. H51 DNA partially digested with Sau3A was cloned
into a cosmid vector (Clare Lister and C Dean unpublished)
that also carried *Agrobacterium* LB and RB T-DNA sequences and
a 35S-NPTII-ocs3' plant selectable marker (CLD04541). 60000
clones were picked into 384 well microtitre plates. In theory

20   there should have been more than a ten-fold redundancy in
this library. The library was hybridized with yeast
artificial chromosome clones (YACs) CIC11A2, EW19A10, EW
12E9, EW 13F8, EW 10A8 and bacterial artificial chromosome
clones (BACs) IGF 5I10, IGF 6N23, TAMU18A10 and IGF2N01

25   covering the *FRI* region (Fig. 1). 93 positively hybridizing
cosmid clones were identified. 59 of these clones were
confirmed by fingerprinting and Southern blots to be unique
and positioned between markers mi51 and CC27P11. The 59
cosmids were than placed into"bins" by hybridization with

30   overlapping BAC clones TAMU8M24, TAMU29E23 and TAMU 10K22.
In addition, the sequence of this genomic region started to
become available as our analysis was proceeding. We then
used the Columbia (Col) genomic sequence as a scaffold to
assemble the H51 cosmid clones into contigs (eg Fig 2). All

35   59 cosmid clones were anchored onto the BAC sequence by end
sequencing the cosmid insert and comparing it to the Columbia
sequence. This analysis determined the orientation of the

33

BACs relative to each other and relative to the chromosome. The colinearity of the H51 line and Col was investigated by comparison of restriction fragments from the cosmids to the map generated from the BAC sequence. Three cosmid contigs corresponding to BACs F5I10, F6N23 and T18A10 were assembled in this manner.

The cosmid library has been constructed in the SURE™ tet^S E. coli strain (Stratagene) to avoid deletions in unstable clones. One drawback of the SURE E. coli strain is the poor conjugation with Agrobacterium. All 59 clones were transformed into E. coli DH5αF' and then conjugated using triparental mating into Agrobacterium. After each step all clones were verified to be identical to the original clone by restriction fragment analysis.

The cosmids were introduced into early flowering Arabidopsis Li5 plants using the vacuum-infiltration Agrobacterium-transformation procedure (Bechtold et al., 1993). Over 300 transformants were obtained with the aim of 5 independent transformants for each cosmid clone.

As the complementation experiments proceeded, further mapping was undertaken. The markers mi51,122 and 204 were anchored onto the BAC contig by sequencing their inserts. Two new RFLP markers using H51 cosmids 40D10 and 32F5 were also developed. The alignment of H51 cosmids to the Col sequence enabled the identification of polymorphic regions and construction of 8 PCR markers in the FRI region (vertical bars, Fig.1 and Table 1). The SSLP marker, UJ3-4 was sensitive enough to identify one heterozygous recombinant plant in a tissue pool of 10 plants. This allowed the isolation of five new recombinant plants by screening batches of 100 plants with just 20 minipreps and 20 PCRs. The CAPS marker UJ12-13 at the other end is less sensitive and was combined with a Southern blot to ensure the detection of recombinants in the same pools. These two markers made it possible to screen large populations of plants to isolate a lot of new recombinants with very little effort.

PCR markers were also used to fine map the recombination

34

break points in the 7 previously identified recombinants.
This allowed the *FRI* region to be shortened ~ 75 kb on
centromere side, eliminating the BAC 2N01 and a gap in the
contig. 50 kb on the telomere side of the *FRI* locus has also
5   been excluded, thereby reducing the number of cosmid clones
for the complementation experiment to a total of 33.

Initially only the clone 84M13 was found to complement
in the T1 generation (Fig. 2). Transformations with
overlapping clones and subclones were pursued in order to
10 narrow down the *FRI* region below the 21 kb of 84M13 (Fig. 3).
Of the overlapping clones only 41J6 could complement whereas
47L16, 163M11 and 23N15 could not. (For some reason 41J6
seems to have much lower transformation efficiency than
84M13.) Some subclones of 84M13; pJU226, pJU222 and pJU223,
15 do also complement whereas pJU234, pJU225 and pJU235 do not.
The complementing cosmid pJU226 contains two copies of an
EcoRI- fragment in a tandem repeat, but it is likely that at
least most of the *FRI* gene is contained within a 5.5 kb XbaI
-EcoRI fragment in the middle of 84M13.

20

Example 2 - The sequence of FRIGIDA.

The H51 genomic sequence of the first 17 kb of 84M13 has
been determined (FRI sequence in Fig.4). RT-PCR experiments
25 have been used to delimit the transcribed region using
primers described in Table 2. Products were obtained using
the primers at the 5' end CLFRI10 -UJ43 but not when CLFRI9 -
UJ43 were used. This predicts the start of transcription to
be between CLFRI-9 and CLFRI10 and there is a putative TATA
30 box 20bp downstream of CLFRI9. The primers CLFRI11, CLFRI 1,
CLFRI2, CLFRI3 all gave products with UJ43 that were the same
size from cDNA made from polyA RNA isolated from H51 plants
as from plasmid DNA indicating that there were no introns in
this region. This then places the beginning of the open
35 reading frame downstream of a stop codon at position 455
(Fig. 4). The first MET codon is at position 574 (Fig 4)
giving a 5' untranslated leader of >360 nucleotides. By using

RT-PCR we were able to verify that this gene is transcribed
at least in leaves of both Li-5 and H51. Furthermore we were
able to amplify the cDNA with primers UJH-UJ37 and to verify
the positions of the first two introns by digestions of RT-

5   PCR products (the second splice site creating a *DraI*
restriction site). The next stop-codon in frame (2880, Fig.
4) is likely to terminate translation. This stop-codon is
well before the EcoRI site (15516) which defines the end of
the complementing *FRI* gene in pJU226. RT-PCR analysis of the

10   3' end of the transcribed region yielded a product between
UJ30 and an oligo dT primer of ~540bp putting the
polyadenylation site around the primer CLFRI16. This would
give a 3' untranslated region of ~500 nucleotides. The most
likely transcribed sequence is shown in Fig. 5, the exact

15   delimitation at both the 5' and 3' ends has still not been
determined. The predicted amino acid sequence from this cDNA
is shown in Fig. 6.

Loss of function FRI *alleles.*

20   An EMS mutagenesis of H51 was undertaken to identify
loss of function alleles of *FRI*. All 40 early flowering
mutants identified were crossed to Li5 to test for allelism
with FRI. None however were allelic.
         In a separate study, about 70 early flowering mutants

25   were isolated from another late ecotype, Sf-2, by fast
neutron radiation. Allelism tests suggested four might be
allelic to *FRI*. A preliminary study of hybridizing BAC DNA
onto DNA isolated from the early lines suggested that 5 of
these FN-mutant lines had deletions in a 15 kb region within

30   the cosmid clones 84M13 and 41J6. We have been able to
complement 3 of these mutants: FN13, FN233 and FN235 with
84M13, thereby confirming that the mutations causing the
early flowering are indeed in this region. The FN mutants
were analyzed by Southerns. FN233 and FN235 were confirmed

35   to have large deletions in the region corresponding to 84M13,
resulting in lines with intact promoter regions but lacking
most of the coding part of the *FRI* gene (Fig.3). FN13 could

be interpreted as carrying a recombination or a large
insertion at the very beginning of *FRI* (Fig.3). The analysis
of the FN mutants and the subclones of 84M13 support the
notion that dominant late flowering alleles of *FRI* are
5   expressing an active protein whereas the early lines have
non-functional alleles of *FRI*.

Allelic variation at FRIGIDA.

10   When the sequence of the H51 *FRI* allele is compared to
that in Col there are 6 nucleotide substitutions in the *FRI*
region. The first two (positions 398 and 432) are in the
5'UTR. Two result in amino acid changes (1010 G in H51, A in
Col) and 1017 (G in H51, A in Col) changing Gly to Glu and
15   Met to Ile. A fifth substitution is found in the middle of
the first intron and the last is a synonymous substitution at
position 1971. The largest difference between the H51 and
the Col sequence is a 16 bp deletion in Col (1509-1524).
This deletion is in the first exon of *FRI* just before the
20   donor splice and would result in a change of reading frame in
Col relative to H51. The frame shift would lead to
translational termination at a stop codon in the next exon
and synthesis of a protein about half the size of the
predicted H51 allele.
25   In order to characterize the *FRI* alleles of early and
late ecotypes two PCR markers were constructed. These
markers allowed us to investigate if the Gly polymorphism
(generation of a BsmFI site), or the 16 bp deletion occurred
in other ecotypes (Table 3). The late ecotypes fell into two
30   classes and either showed the same pattern as H51, or
contained the Gly polymorphism. The early flowering ecotypes
fell into two classes based on the use of these markers. One
class of early ecotypes looked like Col, ie. they had both
the changes relative to H51. Another class of plants looked
35   like H51 with the two markers, but consisted of early and
intermediate ecotypes. The early FRI allele from one member
of this class, Ler, was sequenced to find out if there were

other mutations in the gene. A deletion of 375 bp and
insertion of a novel 31bp was found at the 5' end of the open
reading frame of the Ler *FRI* allele. This deletion would
abolish the likely translation initiation codon, thereby

5   probably creating a null allele of *FRI*. A perfect match to
the first 19 bp of these 31 bp is found 64 bp 3' of the
insertion. 9 bp of the 19 bp repeat is also found in the
second half (12 bp) of the 31 bp insertion. It is quite
plausible that the insertion was created at the same time as

10   the deletion arose and tells something about how the deletion
occurred and/or how the DNA lesion was repaired. The 31 bp
insertion contains an ATG codon so translation may initiate
but would result in a short peptide in a different frame to
FRI. There were only two other differences between the Ler

15   sequence and that from H51. One is the same substitution
that is found in an intron in the Col sequence. The other
change would have resulted in an amino acid change if the
promoter region of Ler was intact. However, this is a
conserved substitution of Leu to Ile. It is plausible that

20   both these changes took place before the deletion inactivated
the gene.

When the ecotypes were re-examined in the promoter
region with PCR, this mutation was also found in the ecotypes
Dijon and Gr. The finding of the promoter deletion in other

25   ecotypes beside Ler, argues that it was present in the Ler
ancestor and is not the result of any recent mutagenesis of
the Ler parent. It is possible that Dijon and Gr contain a
late *FLC* allele unlike Ler and that this difference is
responsible for their intermediate flowering time.

30   A third group of early flowering ecotypes do not
contain either the 16bp deletion present in Col or the larger
deletion present in Ler.

It is possible that representatives from this class will
contain further polymorphisms. From the results above it

35   appears that there are at least two different *FRI* alleles
that confer late flowering. There are at least two different
alleles of *FRI* that result in early flowering, that appear to

be *FRI* loss of function alleles and that have arisen
independently of each other from a late allele of *FRI*. It is
interesting to note that the ancestor of Arabidopsis was
probably late flowering, carrying a functional allele of *FRI*.
5   This makes it likely that an orthologue of *FRI* will be found
in related species and might also form the genetic basis for
variation in flowering time among them in the same manner it
does in Arabidopsis.

10   Example 3 - Predicting FRIGIDA function.

Database predictions show FRI to be an alpha helical protein
with no membrane spanning domains. Database searches of the
protein sequence against the Arabidopsis protein set revealed
15   two related predicted proteins from elsewhere in the
Arabidopsis genome. One is on BAC TAMU21B4 on chromosome 5
and is annotated as having similarity to IP3 receptor. It
shares 26% amino acid identity and 51% similarity (over 297
amino acids). However when the sequence from TAMU21B4 was
20   run against all available databases using BLASTP no protein
with homology to IP3 was found so the basis of the annotation
is unclear. The second protein mapping to the ESSA I No2
fragment on chromosome 4 is annotated as being a
hydroxyproline rich glycoprotein homologue. It shares 21%
25   amino acid identity and 42% similarity over 236 amino acids.
FRI however does not contain the high proportion of proline
residues characteristic of this family of proteins. These
proteins are unlikely to represent FRI functional homologues.

30   A TBLASTN analysis against the Arabidopsis Genbank
dataset gave the same proteins from the Arabidopsis genome
sequence as described above. In addition a few EST sequences
were detected with very low homology scores. For example EST
clone 3D1T7P showed the low score of p=0.00031 and shared 29%
35   amino acid identity over only 107 amino acids. No EST
corresponding to the FRI Col allele is present in the
databases. A BLASTP analysis against the non-redundant

SwissProt database showed a range of functionally unrelated
proteins all with very low homology scores (e>0.68). A
proportion of the proteins were however related in some way
to the cytoskeleton: yeast intracellular protein transport
5   protein USO1, NUF1 spindle body spacer protein, troponin T
etc. This precipitated us to look more closely at the
secondary structure of FRI. Using programmes available on
the Internet, provided by Andrei Lupas, FRI is predicted to
contain two coiled coil domains (between amino acids 45-100
10   and 400-450). Coiled coil domains are common in cytoskeletal
proteins as they cause the protein to have a rod-like
configuration. They are a hall mark of protein: protein
interactions. Proteins either homo or heterodimerize through
the coiled coil regions (Cohen and Parry, 1986; Lupas et al.,
15   1991).

Example 4 - FRIGIDA *homologues.*

The database searches had revealed the existence of *FRI*
20   homologues showing relatively weak homology (in the order of
26% amino acid identity) to *FRI*. However, the presence of
cross-hybridizing fragments in addition to *FRI* on genomic
Southerns that had been washed at high stringency suggests
that there is at least one closely related *FRI* homologue in
25   Arabidopsis not yet represented in the databases.
Hybridization of a genomic fragment (which also contains an
overlapping superoxide dismutase homologue transcribed on the
complementary strand) to the yeast artificial chromosome
clones that constitute the Arabidopsis physical map shows
30   that a potential *FRI* homologue maps to YAC clones CIC1F8 and
4H4.

Example 5 - production of late-flowering transgenic potato

35   Premature bolting in potato diverts the partitioning of
assimilates from the tuber towards have an extended
vegetative phase are desirable over those which do not.

40

This significantly reduces tuber yield. Accordingly, plants
which have an extended vegetative phase are desirable over
those which do not.

5   An FRI genomic clone, or FRI sequence operably linked to the
    CaMV 35S promoter (Odell et al, 1985 Nature 313, 810-812) is
    introduced into potato plants according to the method of
    Spychalla and Bevan (1993) Plant Tissue Culture Manual B11 1-
    9. Plants are observed having an extended vegetative phase.

10 As an alternative, an additional construct comprising VRN2 in
    antisense operably linked to the CaMV 35S promoter is also
    introduced into a plant. A decreased vernalisation response,
    and delayed flowering time results.

15 <u>TABLES</u>

    <u>Table 1</u>

20 | <u>MARKER</u> | <u>CLOSE RFLP POSITION</u> | <u>BAC</u> | <u>PRIMER</u> | <u>SEQUENCE</u> |
    |---|---|---|---|---|
    | UJ3-4 | CC27P11 | F2N01 98105-100090 | UJ3 UJ4 | GTC GGA CCA CAG TTG ATA AGA AT
TCG CAG ATA AGG AGA CTA ACC A |
    | UJ8-9 | 40D10 | T18A10 47043-472167 | UJ8 UJ9 | GAG TTC CGC GAC CCT TTA C
TAG TTT CCG TTG ATA TGT GAT TT |
    | UJ10-11 | mi122 | T18A10 90704-91517 | UJ10 UJ11 | TAA GAA GCC GAA AAC AAA AGG AT
AGG GTA AAA ACT GCA GAT GAA AAT |
    | UJ12-13 | mi51 | F5I10 105629-106219 | UJ12 UJ13 | CGG GGT CAG GTA ATA GCA CAC
GGT TTT CGG ATT TCG GAT TTT A |
    | UJ14-15 | mi51-mi204 | F5I10 54494-55506 | UJ14 UJ15 | AAT TCA ACC GCA TCG TAT CAG
TAT CAG CCG TAT CAA CCA CAT T |

UJ18-19	mi51-mi204	F5I10	UJ18	CCA CCG TTA GTC TAT GCC TGA GTA
		49820-50835	UJ19	GAT GGG TCG GTG GGT GAA C
UJ20-2	mi204-mi122	F6N23	UJ20	ACC GCA GAA GCA GCA TTA GC
		52046-53173	UJ21	CTC CGC GCA GGT GAT TTG
UJ24-25	mi204-mi122	F6N23	UJ24	CTC CCG ACA GTT TCT TTG ACG
		45586-47168	UJ25	CCT GTT CCT GGC GGT GTA G

Table 1 (continued)

	MARKER	ENZYME	Ler	H51	Li5	Col	Sf2 FRI (Col)
	UJ3-4	SSLP	Col	1400	Col	1986	Col
5	UJ8-9	SSLP	1300(+1 200)	1300(+15 00)	Col	2125(+15 00)	Col
	UJ10-11	HindIII	H51	814	Col	587+227	H51
	UJ12-13	Xbal	H51	430+160	Col	590	H51
	UJ14-15	BspHI	H51	520+490	Col	1013	
	UJ18-19	Sall	H51	1016	Col	653+363	H51
10	UJ20-21	HindIII	H51	650+310+ 140	Col	650+450	H51
	UJ24-25	SSLP	Col	1400	Col	1583	H51

Table 2

OLIGO NAME	OLIGO SEQUENCE	POSITION IN GENOMIC SEQUENCE (5' > 3')
UJF	AGT ACT CAC AAG TCA CAA C	1 > 19
CLFRI-8	GGG ATT ATC GTG TTT GAA G	49 > 67
CLFRI-9	CAT ATT ACC GAG CAA GAA C	130 > 149
UJO	CAG TGG TTT ATA ACA TGT C	183 > 165
CLFRI-10	CAT GTC GTA ATC ATG CAA C	213 > 231
CLFRI-11	GTG CGT AGA TTC AAT TAT TTG	276 > 296
CLFRI-1 -	CAA ATA CAT ATT TTC ATA AGC	349 > 370

```
      UJG          CTA AAC ATA TAA CGA TTA CC           386 > 405

      UJ41         CGT TTT CTC CTA ATT AAA AG

   5  UJ41-2       CGT TTT CTC CTA CTT AAA AG           420 > 439

      CLFRI-2      CTT CAC AAT ATA CAG TTC A            477 > 495

      CLFRI-3      GTG GAA ATT AGG GCT TCT G            529 > 547

  10  RI-CLFRI-3 CCA GAATTC GTG GAA ATT AGG GCT TCT G   529 > 547

      UJP          GTG GAT AAT TGG ACA TGA G            589 > 571

  15  UJH          CCA TAG ACG AAT TAG CTG C            746 > 764

      UJ43         GAA GAT CAT CGA ATT GGC              801 > 794

      UJ32         GGT TTA TTC GAC GTC TCC              1001 > 984

  20
      UJQ          GCT TTG AAA TTG GCC AAG G            1106 > 1123

      UJ33         AGA CTC CAG TAT AAG AAG              1242 > 1225

  25  UJ26         AGA TTT GCT GGA TTT GAT AAG G        1440 > 1461

      UJ34         ATA TTT GAT GTG CTC TCC              1664 > 1647

      UJ35         CTC AAA TGA CTC CTT GCT C            2058 > 2040

  30
      UJ28         TGC GAA AGA ACT ACC AGG ATG          2258 > 2278

      CLFRI-6      CAG CTC TTG TGA GTA GTT AC           2527 > 2546

  35  UJ29         ATT CAT ACT CTC CAG GTC A            2662 > 2680

      UJ37R        AAC AAC AGT TAC CAT ATG G            2767 > 2785
```

44

UJ37   ACC ATA TGG TAA CTG TTG   2786 > 2769

UJ30   TTA TCC AAT CAA AGG TCT CC   2835 > 2854

5   CLFRI-13   GTC ATT TAT TTA ACT CCC AA   2932 > 2951

CLFRI-13R-RI CGC GAATTC TTG GGA GTT AAA TAA ATG AC   2951 > 2932

CLFRI-14   GCT CCT GTA ATT GAC ATT TAA G   3003 > 3024

10   CLFRI-15   CAC TAT CTA AAT AGA CCT C   3077 > 3095

UJ44   TGC GGA TTC CAA CCT TG   3171 > 3187

15   CLFRI-12   GAT TGT CAA GCT CAA GTT GG   3298 > 3279

UJ38   CAA GAT CAA AGA CTG CTA AAT C   3360 > 3339

CLFRI-16   GTG AGT GTA TCT AGT GTT CA   3391 > 3372

20   UJ39   CAG AAG CCT CCG GCG AAC   3761 > 3744

GENERAL METHODS

Growth conditions and measurement of flowering time

The majority of the flowering time analyses were
undertaken in a greenhouse under UK summer conditions. Some
plants were also grown under defined conditions in Sanyo
Gallenkamp Controlled Environment rooms at 20°C. The 16h
photoperiod comprised of 10 hours of 400 Watt metal halide
power star lamps supplemented with 100 watt tungsten halide
lamps and 6 hours using only the tungsten halide lamps. The
combination of lamps used for the 10 hour period provided a
PAR of 92.9 µmoles photons m-2 s-1 and a red:far red ratio of
1.49. The 8 hour extension produced PAR of 14.27 µmoles m-2
s-1 and a red:far-red ratio of 0.66.

Flowering time was measured by counting the total number
of leaves, excluding the cotyledons. The close correlation
between leaf number and flowering time was previously
demonstrated for Landsberg erecta and *fca* alleles (Koornneef
et al., 1991).

Cosmid and RFLP markers.

DNA of pUC clones mi204, mi51, mi122 were obtained from
Bob Whittier and used according to (Lui et al., 1996).
Cosmids g8802 and g3843 were obtained from Howard Goodman
(MGH, Boston), cultured in the presence of 50 mg/l kanamycin,
and maintained as glycerol stocks at - 70°C. CC27P11 was
isolated as a random cosmid clone carrying Columbia genomic
DNA and mapped by hybridization to yeast artificial
chromosome clones (Schmidt et al., 1996). Total cosmid DNA
was used as radiolabelled probe to YAC and BAC library colony
filters and plant genomic DNA blots.

DNA extractions

Arabidopsis DNA was performed by a CTAB extraction
method described by (Dean et al., 1992).

RFLP analysis.

46

Two to three micrograms of plant genomic DNA was
prepared from the parental plants used in the crosses and
cleaved in a 300 µl volume. The digested DNA was ethanol
precipitated and separated on 0.7% agarose gels and blotted

5   onto Hybond-N filters. Gel transfer to Hybond-N,
hybridisation and washing conditions were according to the
manufacturer's instructions, except that DNA was fixed to the
filters by UV Stratalinker treatment and/or baked at 80°C for
2 h. Radiolabelled DNA was prepared by random hexamer

10   labelling. Radiolabelled probe DNA was hybridised to the
filters to identify RFLPs.

RNA extractions

RNA was extracted using a method described by (Dean et
15   al., 1985). polyA RNA was isolated using the polyAtract[R] mRNA
isolation system (Promega).

RT-PCR analysis

Total RNA was isolated from whole seedlings at the 2-3
20   leaf stage growing under long days in the greenhouse. For
first strand cDNA synthesis, 3 µg of RNA in a volume of 12 µl
was heated to 70°C for 3 minutes, and then quickly cooled on
ice. 7.5 µl of reaction mix was made containing 1 µl of
RNAsin, 1 µl of primer (20µM), 4µl of 5x reverse

25   transcriptase buffer (250mM TrisHCl pH8.3, 375mM KCl, 15mM
MgCl2), 2µl DTT (100mM), 1µl dNTP (10mM), 1µl reverse
transcriptase (200 units, M-MLV Gibco). This reaction mix
was then added to the RNA creating a final volume of 20µl.
The mixture was incubated at 37°C for 2 hours and then

30   diluted to 100 µl with water.
5µl of the diluted first strand synthesis reaction was
added to 95µl of PCR mix containing 16µl 1.25mM dNTP, 10µl
10xPCR buffer (Boehringer plus Mg), 5µl of a 20mM solution of
each of the primers, 63.6µl of water and 0.4µl of 5 units/µl

35   Taq polymerase (Boehringer or Cetus Amplitaq). The reaction
was performed at 94°C for 1 minute, 34 cycles of 55°C for 1
minute, 72°C for 2 minutes and then finally at 72°C for 10

minutes.

DNA sequencing

DNA was sequenced using cycle sequencing with Big Dye
5   terminators (PE Applied Biosystems) and then run on an ABI
377 sequencing machine.

Preparation of H51 cosmid library

10   Plant DNA preparation. 30g of H51 plant tissue (in samples of
3g) was ground to a fine powder with liquid nitrogen using a
pestle and mortar. The powder was transferred to a 50ml
centrifuge tube and 20ml of urea extraction buffer [Urea
extraction buffer; 8.0M Urea (reagent grade), 0.35M NaCl,
15   0.05M Tris pH7.5, 0.02M EDTA, 2% sarcosine, 5% phenol (added
after autoclaving)was added]. The tubes were inverted gently
about 30 times and 0.8ml of 20% SDS was added. The tubes
were incubated at 65°C for 10minutes. 20ml of
phenol/chloroform/IAA was added and the tubes inverted 30
20   times as before and then spun at 3,000 rpm for 20 minutes.
The phenol/chloroform/IAA extraction was repeated and the
aqueous phase removed to 4ml of 5M potassium acetate, 16ml of
isopropanol was added and gently mixed. The nucleic acid was
pelleted immediatelty by centrifugation at 3000rpm for 20
25   minutes. The resulting pellet was dissolved in 4.1ml of TE.
4.0g of caesium chloride and 0.1ml of 5mg/ml of ethidium
bromide was added. The resulting liquid was transferred to
an ultracentrifuge tube and spun overnight in a vertical
rotor at 53,000rpm. The genomic DNA was removed using a
30   syringe and wide-bore needle into a 15ml Falcon tube.
An equal volume of butanol saturated with sodium
chloride was added to the sample and mixed gently, the
aqueous phase was removed to a clean tube and the process
repeated until the aqueous phase appeared colourless.
35   The caesium chloride was removed by dialysing the sample
against 21 of TE for 4 hours and then overnight in 21 of TE.
The clean DNA was precipitated in a 30 ml Corex tube

with 0.1 volume of 3M Sodium acetate and 2 volumes of 100%
ethanol overnight at 4°C. The Corex tube was spun for 10
minutes at 3000rpm, the ethanol poured off and the pellet
dried under vacuum, it was re-dissolved in 500ul of TE
5   overnight at 4°C. The OD^{260} of the sample was then measured
to ascertain the amount of DNA.

Preparation of pure 04541 vector DNA 50ml of LB +
tetracycline (10mg/ml) was inoculated with a single colony of
10   E.coli carrying the 04541 vector (Bancroft et al., 1997).
This was grown overnight at 37°C and plasmid DNA isolated
using the procedure described in (Birnboim and Doly, 1979).
The DNA was purified using caesium chloride as described
above.

15

Partial digestion of genomic DNA A small-scale test partial
was performed on 10mg of genomic DNA using the restriction
endonuclease Sau 3A. This was done by using different
amounts of the enzyme from 1 unit of enzyme/ug DNA in a
20   reaction mix to 0.08 units of enzyme/mg DNA in a reaction
mix. The digest mixes were incubated for 1 hour at 37°C, the
reactions were stopped by the addition of 1ml of 0.5M EDTA.
The digests were run on an 0.4% agarose gel overnight with
DNA size markers which have fragments between 9kb and 30kb.
25   The gel was stained with ethdium bromide solution and a
polaroid photograph was taken. The lane (and therefore the
most suitable enzyme concentration) containing the most
amount of DNA of the required size (between 15kb and 25kb)
was noted.
30   Having carried out the small scale test, the rest of the
genomic DNA was used in a large scale digestion. The
conditions were exactly the same as before except half the
enzyme concentration as determined in the small scale test
was used. This gave the maximum amount of fragments of the
35   required size. After incubation, 1mg aliquots were removed
from the reaction mixes and run on a 0.4% agarose gel
overnight to confirm that there were fragments of the

expected size. The rest of the partially digested DNA was
precipited with 0.1 volumes of sodium acetate and 2 volumes
of 100% ethanol. The DNA precipitate was rinsed in 70%
ethanol and re-dissolve in 270ml of water.

Treatment with calf intestinal alkaline phophatase (CIP) 30ml
of CIP buffer and 2.5 units of CIP were added to the 270ml of
the partially digested genomic DNA. The mix was incubated at
37°C for 20 minutes and then at 68°C for 15 minutes. 30ml of
sodium acetate and an equal volume of phenol/chloroform/IAA
were added, mixed and spun for 5 minutes at 13,000rpm. The
aqueous phase was removed to 2 volumes of ethanol and
precipitated as above. The partially digested, CIP treated
DNA was rinsed in 70% ethanol, dried and re-dissolved in
500ml of TE.

Size fractionation of the genomic DNA A sucrose gradient was
prepared with 10-40% sucrose solutions. The DNA in TE was
loaded onto the top of the gradient and it was centrifuged
for 18 hours at 20,000rpm in a SW41 rotor. 20 fractions of
300-500ml were collected from the bottom of the tube and 10ul
of each fraction was run overnight on an 0.4% agarose gel.

Purification of fractions of the right size. The fractions
from the sucrose gradient which contained partially digested
DNA of between 15kb and 20kb were diluted 1 in 4 with water.
The DNA was precipitated by the addition of 1/30 volume of 3M
sodium acetate pH5.5, 1ul tRNA, and 0.7 volumes of
isopropanol and incubation on ice for 60 minutes. The
precipitate was pelleted by centrifugation for 20 minutes at
13,000rpm, it was rinsed in 70% ethanol and dried under
vacuum. The fractions were redissolved in 5ul of water and
then pooled into one tube.

Preparation of vector DNA 10mg of 04541 was digested with
BamH1. An aliquot of 100ng was removed and run on an 0.5%
agarose gel with an aliquot of uncut vector to confirm that

50

the vector had been digested completely. The rest of the
digest was extracted with an equal volume of
phenol/chloroform/IAA and precipitated with 0.1 volumes of
sodium acetate and 2 volumes of 100% ethanol washed in70%
5   ethanol and dried under vacuum.

Ligation of the digested vector and the purified fractions
1mg of purified insert DNA was combined with 0.25mg of
digetsted vector DNA in a 7.5ml volume. 0.5ml was removed
10 and stored at -20°C (unligated control). The remaining
reaction was incubated at 65°C for 5 minutes, then for 20
minutes at 42°C and finally for 2 hours at room temperature.
1ml of Tris-HCl/MgCl2,   0.5ml of 0.1M ATP, 0.5ml of 5:1000
dilution of b-mercaptoethanol and 1.0ml of T4 DNA ligase were
15 added and the mix was incubated for 18 hours at 12°C. The
unligated control and 0.5ml of ligation were run on a 0.5%
minigel to check ligation.

Preparation of plating cells SURE^{TM}TetS (Stratagene) were
20 streaked out on L agar plate and grown overnight at 37°C. A
single colony was used to inoculate 50ml of LB maltose plus
0.5ml 1M MgSO4. This was shaken for 4-6 hours at 37°C until
the OD600 was between 0.5 and 1.0. The bacteria were
pelleted at 2,000rpm for 10 minutes and resuspended in 20ml
25 of sterile 10mM MgSO4. The OD600 was measured and more 10mM
MgSO4 was added until the OD600 reached 0.5. The cells were
stored at 4°C.

Packaging of ligation The packaging into the phage was
30 performed using Gigapack II XL packaging extracts from
Stratagene and the protocol accompanying the extracts was
followed. 10ml of phage stock was added to 100ml of phage
buffer and 200ml of plating cells and incubated at 37°C for
20 minutes. 1ml of L broth was added and incubation
35 continued for another hour. 0.1ml and 0.5ml aliquots were
spread on a L agar tet10mg/ml plate and incubated overnight.
Colonies were counted and the titre was calculated.

Transformation of Arabidopsis

The cosmids containing DNA from the vicinity of *FRI* were mobilised into Agrobacterium tumefaciens C58C1 using triparental mating (Hoekema et al., 1983). The T-DNA was
5   introduced into Arabidopsis plants using the vacuum-infiltration technique as described in (Bechtold et al., 1993).

10 *References*

Bancroft et al. (1997). Weeds World 4ii,

Bechtold, N., Ellis, J. and Pelletier, G. (1993). *In planta*
15 *Agrobacterium* mediated gene transfer by infiltration of adult
Arabidopsis thaliana plants.   C.R.Acad.Sci.Paris 316,1194-1199.

Birnboim, H.C. and Doly, J. (1979). Nucleic Acids Res.
20 7(6),1513-1523.

Burn, J.E., Smyth, D.R., Peacock, W.J. and Dennis, E.S.
(1993). Genes conferring late flowering in *Arabidopsis thaliana*. Genetica 90,145-157.
25
Clarke, J.H. and Dean, C. (1994). Mapping *FRI*, a locus controlling flowering time and vernalization response in *Arabidopsis thaliana*. Mol.Gen.Genet. 242,81-89.

30 Cohen, C. and Parry, D.A.D. (1986). a-helical coiled coils – a widespread motif in proteins.   TIBS 11,245-248.

Dean et al. (1985).   Embo J. 4,3055-3061.

35 Dean et al (1992).   Plant J. 2,69-81.

Hoekema et al (1983).   Nature. 303,179-180.

Koornneef et al (1991). Mol.Gen.Genet. 299,57-66.

Koornneef, M., Blankestijn-de Vries, H., Hanhart, C., Soppe,
W. and Peeters, T. (1994). The phenotype of some late-
flowering mutants is enhanced by a locus on chromosome 5 that
is not effective in the Landsberg *erecta* wild-type. Plant J.
6, 911-919.

Koornneef, M., Alonso-Blanco, C., Peeters, A.J.M. and Soppe,
W. (1998). Genetic control of flowering time in Arabidopsis.
Annu.Rev.Pl.Phys.Pl.Mol.Biol. 49, 345-370.

Lee, I., Bleecker, A. and Amasino, R. (1993). Analysis of
naturally occurring late flowering in *Arabidopsis thaliana*.
Mol. Gen. Genet. 237, 171-176.

Lee, I., Michaels, S.D., Masshardt, A.S. and Amasino, R.M.
(1994). The late-flowering phenotype of *FRIGIDA* and mutations
in *LUMINIDEPENDENS* is suppressed in the Landsberg *erecta*
strain of *Arabidopsis*. Plant J. 6, 903-909.

Lui, Y-G., Mitsukawa, N., Lister, C ., Dean, C. and Whittier,
R.F. (1996). Isolation and mapping of a new set of 129 RFLP
markers in *Arabidopsis thaliana* using recombinant inbred
lines. Plant J. 10, 733-736

Lupas, A., van Dyke, M. and Stock, J. (1991). Predicting
coiled coils from protein sequences. Science 252, 1162-1164.

Napp-Zinn, K. (1957). Utersuchungen zur genetik des
kaltebedurfnisses bei *Arabidopsis thaliana*. Zeitschrift fur
inducktive Abstammungs - und Vererbungslehre 88, 253-258.

Napp-Zinn, K. (1962). Úber die genetischen Grundlagen des
VernalisationsbedÝrfnisses bei *Arabidopsis thaliana*. I. Die
Zahl der beteiligten Faktoren. Z.Vererbungsl. 93, 154-163.

Napp-Zinn, K. Theory of vernalization - new experiments with *Arabidopsis*. In: *Ist International Symposium on Arabidiopsis Research*, Gšttingen: Suppliment, Arabidopsis Information Service, 1965, p. 56-61.

Napp-Zinn, K. Vernalization - environment and genetic regulation. In: *Manipulation of Flowering*, edited by Atherton, J.G. London: Butterworths, 1987, p. 123-132.

Osborn, T.C., Kole, C., Parkin, I.A.P., Sharpe, A.G., Lydiate, D.J. and Trick, M. (1997). Comparison of flowering time in *Brassica rapa*, *B. napus* and *Arabidopsis thaliana*. Genetics 146,1123-1129.

Sanda, S.L. and Amasino, R.M. (1996). Interaction of *FLC* and late-flowering mutations in *Arabidopsis thaliana*. Mol.Gen.Genet. 251, 69-74.

Schmidt, R., West, J., Love, K., et al. (1995). Physical map and organization of *Arabidopsis thaliana* chromosome 4. Science 270, 480-483.

Schmidt et al (1996). Plant J. 9,755-765.

SEQUENCES OF VRN2

Landsberg erecta VRN2 cDNA (SEQ ID NO: 57)

CAAGCTTCTTCAATTTTGCTTGCTCTCTCTTACACAGCAATCGGTGTTTTCGCAGCTTTCA
GGCCTCAATCCAAGACATTCTATATAAGCATATTGCAGAAGAGGCGGTTCTAATTGTTGCAT
TGAGTTTATCGCTATGACGTAGGGAAATTCTAATTAGGGGAGGCCTCAGAGTTTGCACTAA
CTTCATAATCGGCTCTTGACGTTGTTGAGTGTAATTGAACAAGAATGTGTAGGCAGAATTGT
CGCGCGAAATCCTCACCGGAGGAAGTGATTTAACTGATGAGAATCTCTTGATATATTGTAA
ACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTAGGCAACCCATCGTTTCTTCCAA
GATGCTTGAACTACAAAATTGGAGCAAGCGCAAAAGAAAGTCAAGATCTACTGGGATGGTA
GTTTTCAACTATAAGGATTGTAATAACACATTACAGAAAACTGAAGTTAGGGAGGATTGTTC
TTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGGGGCTGCAATTTCATTTGAATTCAT

CTCATGATTTATTTGAATTTGAGTTCAAGCTTTTCGAAGAATACCAGACAGTTAATGTTTCT

GTAAAACTTAATTCCTTCATATTTGAGGAAGAAGGAAGTGATGACGATAAATTTGAGCCCTT

CTCTCTCTGCTCGAAACCTCGTAAGCGGAGACAAAGAGGTGGCAGAAATAACACCAGGAGAC

TTAAAGTATGCTTTTTACCGTTGGATTCACCCAGTTTAACTAATGGCACAGAAAATGGAATC

ACCCTACTTAATGATGGAAACCGTGGTTTAGGATATCCCGAGGCAACAGAGCTTGCTGGACA

AATTTGAGATGACCAGCAACATTCCACCAGCCATAGCCCACTCTTCTCTGGACGCTGGTGCTA

AAGTTATATTGACAAGCGAAGCTGTGGTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAG

CGATCAGAGGCTAGAAGCCACCTACTTCTTCAGAAACGCCAATTCTATCATTCTCACAGAGT

CCAGCCAATGGCGCTTGAGCAAGTAATGTCTGACCGGGATAGCGAGGATGAAGTCGATGACG

ATGTTGCAGATTTTGAAGATCGCCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAA

AAGCAATTCATGCATCTTTGGAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCA

TATCTCTTGGGCATGTGAAGCATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTTACTCAT

CACTCTTCTGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCA

GCCACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCAC

CACCAACAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTG

TGGATCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGA

GACAAAGTAATTAAATAGGAAAATCTCCGGCTTTTATGATACCGATTTATCGGATTGTAACT

TATTCTTCTTTCTTAAAAAATTGTTTAGGAGCAAACAAATTTTTTATATGTTAGTGTATTCA

ACTGATTACATTTTTAGTTAAAAAAAAAAATGGATTCTGCTTATAACT

Columbia VRN2 cDNA (SEQ ID NO:58)

CAAGCTTCTTCAATTTTGCTTGCTCTCTCTCTTACACGGCCAATCGGTGTTTTCGCAGCTTT

CAGGCCTCAATACAAGACATTCTATATAAGCATATTGCAGAAGAGGCGGTTCTAATTGTTGC

ATGGAGTTGAACAATATGACGTAGGGAAATTCTAATTTAGGGGAGGCCTCAGAGTTTGCACT

AACTTCATAATCAGCTCTGGACGTTGTTGATTGTATTTGAACAAGAATGTGTAGGCAGAATT

GTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAGAATCTCTTGATATATTGT

AAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTAGGCAACCCATCGTTTCTGCC

AAGATGCTTGAACTACAAAATTGGGGCAAAGCGCAAAGAAAGTCAAGATCTACTGGGATGG

TAGTTTTCAACTATAAGGATTGTAATAATACATTACAAAGAACTGAAGTTAGGGAGGATTGT

TCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGGGGCTGCAATTTCATTTGAATTC

ATCTCATGATTTATTTGAATTTGAGTTCAAGCTTTTGGAAGAATACCAGACAGTTAATGTTT

CTGTAAAACTTAATTCCTTCATATTTGAGGAAGAAGGAAGTGATGATGATAAATTTGAGCCC

TTCTCTCTGCTCGAAACCTCGTAAGCGTAGACAAAGAGGTGGCAGAAATAACACCAGGAG

ACTTAAAGTATGCTTTTTACCGTTGGATTCACCCAGTTTAGCTAATGGCACAGAAAATGGAA

TTGCCCTGCTGAATGATGGAAACCGTGGTTTAGGATATCCCGAGGCAACAGAGCTTGCTGGA

CAATTTGAGATGACTAGCAACATTCCACCAGCCATAGCCCACTCTTCTCTGGACGCTGGTGC

TAAAGTTATATTAACAACCGAAGCTGTGGTCCCTGCTACTAAGACAAGAAAGTTATCTGCTG

AGCGATCAGAGGCTAGAAGCCACCTACTTCTTCAGAAACGCCAATTCTATCATTCTCACAGA

GTCCAGCCAATGGCGCTTGAGCAAGTAATGTCTGATCGGGATAGCGAGGATGAAGTCGATGA

CGATGTTGCAGATTTTGAAGATCGCCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATG

5   AAAAGCAATTCATGCATCTTTGGAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGT

CATATCTCTTGGGCATGTGAAGTATTTTCAAGATTTTACGAGAAAGAGTTGCACTGTTACTC

ATCACTCTTCTGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACT

CAGCCACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATACCTCAGTCACTAAC

AACAACAACAACAGTGTGGATCATCCCAGTGACTCAAACACCAACAACAATAACATTGTGGA

10 TCATCCGAATGACATAAAAAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAGACA

AGTAATTAAATAGGAAACACTCCGGTTTAGATGATACCGATCTATCGGATTGTAACTTATTC

TTCTTTCTTAAAAAAATTGTTTAGGAGCAAACAAAGATTTTATTTGTTAGTGTATTCAACTG

ATTACATTTTTAGTTAAAAAAATGGATTCTCCTTAATAACT